Iterative Linear Equation Solvers

Elimination: even if $\#\text{nnz} \in O(n)$,
can require space $\Omega(n^2)$ and time $\Omega(n^3)$.

Iterative: only store matrix and a few vectors.
Approach solution rather than compute it exactly.

Want to solve $Ax = b$, A pos def

Observe $Ax = b \iff \alpha A x = \alpha b$
$$\Rightarrow x + (\alpha A - I)x = \alpha b$$
$$\Rightarrow x = (I - \alpha A)x + \alpha b$$

Iterate on this
$$x_0 = 0$$
$$x_t = (I - \alpha A)x_{t-1} + \alpha b$$

Is called Richardson Iteration

Converges if $\|I - \alpha A\| < 1$

A sym $\Rightarrow \|I - \alpha A\| = \max_i |1 - \alpha \lambda_i|$

where $0 < \lambda_1 < \cdots < \lambda_n$

Set $\alpha = \dfrac{2}{\lambda_1 + \lambda_n}$ gives $\lambda_1, \lambda_n = \pm \dfrac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}$

and $\|I - \alpha A\| = 1 - \dfrac{2\lambda_1}{\lambda_1 + \lambda_n}$

If $\alpha < \frac{2}{\lambda_1 + \lambda_n}$, $\|I - \alpha A\| \le 1 - \lambda_1 \alpha$

Convergence: consider $x - x_t$.

$$x - x_t = \left((I - \alpha A)x + \alpha b\right) - \left((I - \alpha A)x_{t-1} + \alpha b\right)$$
$$= (I - \alpha A)(x - x_{t-1})$$
$$= (I - \alpha A)^t (x - x_0)$$
$$= (I - \alpha A)^t x$$

$$\Rightarrow \|x - x_t\| \le \|x\| \cdot \|I - \alpha A\|^t$$
$$= \|x\| \left(1 - \frac{2\lambda_1}{\lambda_1 + \lambda_n}\right)^t$$
$$\le \|x\| e^{-\frac{2\lambda_1 t}{\lambda_1 + \lambda_n}}$$

To get $\frac{\|x - x_t\|}{\|x\|} \le \varepsilon$,

suffices to have $t = \ln(1/\varepsilon) \frac{\lambda_1 + \lambda_n}{2\lambda_1}$

$$= \left(\frac{\lambda_n}{2\lambda_1} + \frac{1}{2}\right) \ln(1/\varepsilon)$$

key term is $\frac{\lambda_n}{\lambda_1} \triangleq \kappa(A)$,

the condition number.

\# iters $\approx \frac{1}{2} \kappa(A) \cdot \ln(1/\varepsilon)$

For Laplacian of ε-expander, orth to $\mathbb{1}$,

$\kappa(L) \approx (1+2\varepsilon)$ so very fast convergence

---

Issue: might not know λ_1, λ_n, so only guess α.

Can not tell if have answer.

Can measure $b - Ax_t = Ax_0 - Ax_t = A(x_0 - x_t)$

$$= A(I - \alpha A)^t x_0 = (I - \alpha A)^t A x_0$$

$$= (I - \alpha A)^t b$$

So, $\|b - Ax_t\| \leq \|I - \alpha A\|^t \|b\|$,

which could be more useful.

But, $\|b - Ax_t\| \leq \varepsilon\|b\|$ does not imply $x \approx x_t$

Only implies $\|x - x_t\| \leq \varepsilon\|x\| \cdot \kappa(A)$

A view through polynomials.

Can write x_t as $p_t(A)b$, for some polynomial p_t

Check: $x_0 = 0$

$x_1 = \alpha b$

$x_2 = (I - \alpha A)\alpha b + \alpha b$

$x_3 = (I - \alpha A)^2 \alpha b + (I - \alpha A)\alpha b + \alpha b$

$$x_t = \sum_{i=0}^{t} (I - \alpha A)^i \alpha b$$

Will see that $p_t(A) \approx A^{-1}$

First, take the limit as $t \to \infty$

$$p_t(A) \to \alpha \sum_{i \geq 0} (I - \alpha A)^i = \alpha\left(I - (I - \alpha A)\right)^{-1}$$
$$= \alpha (\alpha A)^{-1}$$
$$= A^{-1}$$

In general, a poly p gives an ε-accurate solution if

$$\| p(A) \cdot b - x \| \leq \varepsilon \| x \|$$

$$\iff \| p(A) A x - x \| \leq \varepsilon \| x \|$$

$$\Leftarrow \| p(A) A - I \| \leq \varepsilon$$

Now, we can search for better polynomials.

We need that ~~for~~ λ_i eigvals ~~of~~ A,

$$| p(\lambda_i) \lambda_i - 1 | \leq \varepsilon$$

Def $q(x) = p(x) x - 1$. We need $q(0) = 1$,

$$| q(\lambda_i) | \leq \varepsilon$$

Thm[1] For every $t \geq 1$ and $0 < \lambda_{min} \leq \lambda_{max}$

$\exists$ deg $t+1$ poly $q_t(x)$ s.t.

1. $|q^t(x)| \leq \varepsilon$ for $\lambda_{min} \leq x \leq \lambda_{max}$

2. $q^t(0) = 1$

for $$\varepsilon \leq \frac{2}{\left(1+\frac{2}{\sqrt{\kappa}}\right)^t} \leq 2e^{-2t/\sqrt{\kappa}}$$

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}}$$

A quadratic improvement.

Before proving, consider with Laplacians.

As work orthogonal to $\mathbb{1}$, $\lambda_{min} = \lambda_2$.

A degree t polynomial only moves data t steps through a graph. So, should need $t \geq$ diameter,

For path of length n, $K = \frac{\lambda_{max}}{\lambda_{min}} \approx n^2$,

so $\sqrt{K} \approx n$ iterations makes sense.

For expander, $K \approx$ constant

hypercube, $K \approx \log n$

these can be solved quickly.

Prove Thm1 using Chebyshev polynomials.

Def $T_t(x) = \begin{cases} \cos(t \cdot \mathrm{acos}(x)) & |x| \le 1 \\ \cosh(t \cdot \mathrm{acosh}(x)) & |x| \ge 1 \end{cases}$

To see is a polynomial,

def $T_0(x) = 1 \quad T_1(x) = x$

$$T_t(x) = 2x T_{t-1}(x) - T_{t-2}(x)$$

To verify trig identities,

recall $\cos(\theta) = \dfrac{e^{i\theta} + e^{-i\theta}}{2} \qquad \cosh(\theta) = \dfrac{e^{\theta} + e^{-\theta}}{2}$

if $\theta = \mathrm{acosh}(x) \iff \cosh(x) = \theta$,

and $2x T_{t-1}(x) - T_{t-2}(x)$

$$= 2\left(\frac{e^{\theta} + e^{-\theta}}{2}\right)\left(\frac{e^{(t-1)\theta} + e^{-(t-1)\theta}}{2}\right) - \frac{e^{(t-2)\theta} + e^{-(t-2)\theta}}{2}$$

$$= \frac{1}{2}\left[e^{t\theta} + e^{-t\theta} + e^{-(t-2)\theta} + e^{(t-2)\theta} - e^{(t-2)\theta} - e^{-(t-2)\theta}\right]$$

$$= \frac{1}{2}\left[e^{t\theta} + e^{-t\theta}\right] = \cosh(t\theta)$$

For $|x| \le 1$, $|T_t(x)| \le 1$,

because $\exists\, \theta$ s.t. $\cos\theta = x$, and $|\cos(t\theta)| \le 1$

Claim: For $\gamma > 0$, $T_t(1+\gamma) \ge \frac{1}{2}(1+\sqrt{2\gamma})^t$

using $\mathrm{acosh}(x) = \ln(x + \sqrt{x^2-1})$ for $x \ge 1$

Proof of Thm 1

know $T_t(x)$ has degree t

$T_t(x) \in [-1,1]$ for $x \in [-1,1]$

$T_t(x)$ is monotonically increasing for $x > 1$

$T_t(1+\gamma) \ge \frac{1}{2}(1+\sqrt{2\gamma})^t$

Write $q_t(x) = T_t(\ell(x)) / T_t(\ell(0))$

where $\ell(x) = \dfrac{\lambda_{max} + \lambda_{min} - 2x}{\lambda_{max} - \lambda_{min}}$

$\ell(\lambda_{max}) = -1$

$\ell(\lambda_{min}) = 1$

$\ell(0) = 1 + \dfrac{2\lambda_{min}}{\lambda_{max} + \lambda_{min}} \ge 1 + \dfrac{2}{K}$

By def. $q_t(0) = 1$

for $x \in \lambda_{min}, \lambda_{max}$, $\ell(x) \in [-1,1]$ so

$|T_t(\ell(x))| \le 1$, and $|q_t(x)| \le \dfrac{1}{T_t(\ell(0))} \rightarrow$

$\rightarrow \le \dfrac{2}{(1 + 2/\sqrt{K})^t} \le 2e^{-2t/\sqrt{K}}$